**📘 Lecture Notes: Early Stopping to Prevent Overfitting**

**1. Overfitting Reminder**

* Training a model **forever** will eventually lead to **overfitting**.
* Overfitting occurs when the model fits the training data **too well**, including noise.
* **Solution:** Stop training **early**, before overfitting occurs → **Early Stopping**.

**2. Early Stopping: Definition**

* **Early stopping:** Technique to prevent overfitting by stopping training at the right time.
* Goal: Stop **before the model starts memorizing noise**.

**3. Common Early Stopping Methods**

**A) Preset Number of Epochs**

* Train for a fixed number of epochs (e.g., 100).
* **Problems:**
  + No guarantee that minimum loss is reached.
  + May train for unnecessary epochs → wasted computation.
  + Doesn’t prevent overfitting.
* Simple, but naive method.

**B) Stop When Loss Updates Are Small**

* Stop training when the **relative decrease in loss** is below a threshold (e.g., 0.001 or 0.1%).
* **Advantages:**
  + Ensures minimum is reached.
  + Saves computing power by avoiding useless iterations.
* **Disadvantages:**
  + Still can overfit if training continues too long.

**C) Validation Set Strategy (Most Effective)**

* **Idea:** Use validation loss to detect overfitting.
* **Process:**
  1. Track both **training loss** and **validation loss** during training.
  2. Initially, both decrease as the model learns.
  3. At some point, **validation loss starts increasing** → model is overfitting.
  4. **Stop training immediately** → prevent overfitting.
* **Visual:**
  1. Training loss: keeps decreasing or flattens.
  2. Validation loss: decreases first, then rises when overfitting starts.
  3. Divergence point = red flag → stop.

**4. Combining Techniques**

* Best practice: **combine methods** for more safety:
  + Stop if **validation loss starts increasing** OR
  + Stop if **training loss becomes very small** (weights barely change).
* Combines the benefits of monitoring **both generalization and efficiency**.

**5. Summary / Key Takeaways**

* **Early stopping prevents overfitting** by stopping training at the right moment.
* Methods:
  1. Fixed number of epochs → naive, doesn’t prevent overfitting.
  2. Stop when training loss updates are tiny → saves computation, partial solution.
  3. **Validation set monitoring** → best for preventing overfitting.
* **Best practice:** Combine methods for robustness.

✅ **In one sentence:**

* **Early stopping = stopping training before overfitting occurs, using validation loss as a guide and monitoring weight updates.**